Purification and characterisation of a novel DNA methyltransferase, M.Ahdl

Phil Marks, John McGeehan, Geoff Wilson¹, Neil Errington² and Geoff Kneale*

Biophysics Laboratories, Institute of Biomedical and Biomolecular Sciences, University of Portsmouth, Portsmouth PO1 2DT, UK, ¹New England Biolabs Inc., Beverly, MA 01915, USA and ²School of BioSciences, University of Nottingham, Sutton Bonington LE12 5RD, UK

Received February 18, 2003; Revised March 19, 2003; Accepted April 14, 2003

ABSTRACT

We have cloned the M and S genes of the restriction-modification (R-M) system AhdI and have purified the resulting methyltransferase to homogeneity. M.AhdI is found to form a 170 kDa tetrameric enzyme having a subunit stoichiometry M₂S₂ (where the M and S subunits are responsible for methylation and DNA sequence specificity, respectively). Sedimentation equilibrium experiments show that the tetrameric enzyme dissociates to form a heterodimer at low concentration, with $K_d = 2 \mu M$. The intact (tetrameric) enzyme binds specifically to a 30 bp DNA duplex containing the AhdI recognition sequence GACN₅GTC with high affinity ($K_d = 50 nM$), but at low enzyme concentration the DNA binding activity is governed by the dissociation of the tetramer into dimers, leading to a sigmoidal DNA binding curve. In contrast, only non-specific binding is observed if the duplex lacks the recognition sequence. Methylation activity of the purified enzyme was assessed by its ability to prevent the methyl group being donated by the cofactor S-adenosyl methionine (AdoMet).

Type I and type II R-M systems are readily distinguished. Classical type II systems are two-gene systems, with one gene for each activity; a gene for DNA methylation (M), which codes for a monomeric protein, and a quite unrelated gene for restriction (R), which produces a homodimeric enzyme. For both activities, the DNA recognition sequence is identical: a symmetrical (self-complementary) sequence of, typically, six adjacent base pairs. Type II MTases have two structural domains: the DNA sequence specificity is provided by a target recognition domain (TRD) and methylation activity is found on a separate AdoMet binding domain within the same subunit.

In contrast type I systems consist of three genes. Two genes are required for DNA methylation activity (M and S), the subunits combining to form a trimeric MTase (M₂S) (4–6). An additional gene is required for restriction (R), the three genes together coding for the pentameric endonuclease (R₂M₂S) (7–9). Type I systems recognise an asymmetrical DNA sequence (typically ~13–15 bp) in which the two recognition sites are separated by a non-specific ‘spacer’, 6–8 bp long. The specificity of DNA binding is conferred by the two TRDs of the S subunit, each binding a ‘half-site’ within the DNA recognition sequence (10,11). This locates the two recognition sites approximately one helical turn apart on the DNA.

The discovery of the restriction system AhdI suggested the possibility of a new class of R-M system (type 1½) that combines properties of both type I and type II systems (12). As in type II systems, the AhdI endonuclease is encoded by a single gene; on the other hand, AhdI is typical of type I systems in having two genes resembling the M and S genes that in type I systems are required for assembly of the active MTase. Both M and S genes are required for AhdI methylation activity (G. Wilson et al., manuscript in preparation). Moreover, although (like type I systems) the AhdI DNA recognition sequence is bipartite (GACN₅GTC), it has rotational symmetry characteristic of the recognition site of classical type II systems.

There have been no studies reported on the AhdI methyltransferase to date. Here we report the expression and purification of the methyltransferase M.AhdI and characterisation of the purified enzyme in terms of subunit structure, DNA binding affinity and enzyme activity.

INTRODUCTION

In prokaryotes, methylation of specific sequences serves to protect the host DNA from cleavage by restriction enzymes encoded by the host restriction-modification (R-M) system, whilst allowing the cleavage of foreign DNA (1–3). It thus acts as a form of primitive immune system, protecting the bacterium from invasion by bacteriophage. Methylation occurs at either the cytosine C₅ or at the N₄ or N₆ of adenine,
MATERIALS AND METHODS

Expression and purification of M.AhdI

Escherichia coli JM109 (DE3) cells containing either the pET11a plasmid containing the specificity (S) subunit gene or the pMTYB1 plasmid containing the methylation (M) subunit gene were grown at 30°C in 2x YT broth until an OD600 = 0.6 was obtained. The cells were then induced with 1 mM isopropyl ß-D-thiogalactopyranoside and grown for 3 h. The cells were harvested by centrifugation and the pellets stored at −20°C until required.

The cell pellets were resuspended in 0.1 M NaCl, 50 mM Tris–HCl pH 8.2, 5 mM EDTA, 5 mM dithiothreitol (DTT), 1 mM benzamidine, 0.1 mM phenylmethylsulphonyl fluoride and the M and S cell pastes were mixed in a ratio of 1:1 culture of S subunit to 3:1 of M subunit, this ratio having been found in earlier experiments to give the optimum yield of MTase. Lysis was achieved by sonication at 4°C and the insoluble debris was removed by centrifugation at 48 000 g.

The soluble cell lysate was made to 1 M NaCl in a 30% saturated solution of (NH4)2SO4 and the precipitated protein was sedimented by centrifugation at 17 000 g. The protein pellet was solubilised in loading buffer (0.1 M NaCl, 50 mM Tris–HCl pH 8.2, 5 mM EDTA, 1 mM DTT) and subsequently bound to a heparin affinity column. Once a baseline had been reached, a linear 0.1–0.5 M NaCl gradient in loading buffer was used to elute the target protein. A dialysis step back to loading buffer was followed by final purification using a 24 ml Superose 12HR size exclusion column.

The extinction coefficient of the MTase was calculated to be 47 000 M−1 cm−1 for the MS dimer (94 000 M−1 cm−1 for the M2S2 tetramer) by summing the contributions from the aromatic amino acids in the M and S subunits.

End-labelling of DNA

Oligonucleotides containing the sequences for DNA binding analysis of M.AhdI were annealed to form DNA duplex. The DNA concentration for each duplex was then calculated by the summation of the nucleotide extinction coefficient and the DNA concentration for each duplex was then calculated by the absorbance versus radial displacement were taken at a resolution of 0.001 cm and at intervals of 2–4 h. Sedimentation equilibrium was deemed to have been achieved when there was no change in concentration distribution for a period of at least 4 h. This was performed for rotor speeds of 8000, 12 000 and 20 000 r.p.m.

Analytical ultracentrifugation

A Beckman Optima XL-A analytical ultracentrifuge (Beckman-Coulter, Palo Alto, CA) was used for both sedimentation equilibrium and sedimentation velocity experiments. For the sedimentation velocity experiment a standard double sector cell with 12 mm optical path length was used and filled with 400 µl of solution and 410 µl of solvent/buffer in the respective sectors. The cell was loaded into an An-60 analytical rotor and equilibrated to temperature (20°C) before acceleration to 45 000 r.p.m. Upon reaching speed, scans of absorbance at 280 nm versus radial displacement were taken at intervals of 80 s at a resolution of 0.005 cm.

The sedimentation equilibrium experiment was performed in 6-channel cells of 12 mm optical path, using 70 µl of solution at protein concentrations between 0.6 and 19.4 µM. An aliquot of 75 µl of solvent/buffer was loaded into the corresponding channel. The rotor was accelerated and scans of absorbance versus radial displacement were taken at a resolution of 0.001 cm and at intervals of 2–4 h. Sedimentation equilibrium was deemed to have been achieved when there was no change in concentration distribution for a period of at least 4 h. This was performed for rotor speeds of 8000, 12 000 and 20 000 r.p.m.
RESULTS

Sub-cloning, expression and purification

Co-expression of the S and M subunits of MAhdI could not be achieved in a single expression construct (13). For this reason the two subunits were expressed in separate bacterial cell cultures. The S subunit gene was sub-cloned into the pET11a expression vector. Problems were encountered with plasmid stability and cell toxicity, but with careful culturing at 30°C significant soluble expression could be achieved. The M subunit expression clones were much more stable and less toxic to the cell, although it was rather more difficult to produce soluble target protein from them. The M subunit gene was sub-cloned into a series of different expression vectors (pTYB1, pTYB11, pET11a and pET42a). Optimal expression of soluble protein was found using pTYB1 when grown at 30°C. The expression vectors chosen for the M and S subunits both produce unfused proteins, thus avoiding any possible problems caused by additional protein sequences at the N- or C-termini when forming the multisubunit complex (14).

Trial experiments showed that higher yields of MAhdI were produced by mixing crude cell extracts from the M and S expression systems, rather than reconstitution of the enzyme from its purified subunits (13). The bacterial cell pastes for the S and M subunits were mixed before lysis to form the MAhdI enzyme. The subsequent soluble cell extract was fractionated using ammonium sulphate precipitation to remove contaminating bound DNA from the MAhdI. The precipitation was followed by heparin affinity chromatography, which resulted in relatively pure MAhdI but with some free S subunit also present since this elutes at a similar salt concentration. A final step using size exclusion chromatography was required to separate MAhdI from the free S subunit contaminating the peak on the heparin column (Fig. 1). Replication of this sample to the gel filtration column confirmed that it elutes as a single species, and no peak corresponding to free S subunit was observed on subsequent runs.

MAhdI stoichiometry

The purified MAhdI MTase was applied to a calibrated Superose 12 gel filtration column to give an estimate of the size of the multisubunit enzyme. When loaded at 100 μM concentration, the enzyme elutes in a single peak with an apparent molecular mass of 154 kDa with respect to molecular weight markers. When loaded at lower concentrations (5 and 20 μM), the enzyme also elutes as a single peak, but with a slightly lower estimated molecular weight (144–147 kDa). Dynamic light scattering measurements (Fig. 2a) indicated a hydrodynamic radius of 5.4 nm, from which a molecular mass of 175 kDa is estimated by comparison with typical globular proteins. This is in reasonable agreement with the calculated Mr for the heterotetramer M₂S₂ of 170 398 Da, based on the calculated Mr of the S and M subunits (25 422 and 59 777 Da, respectively). However, both these techniques are unable to separate out the effects of size and shape if the molecule is elongated. Static (Rayleigh) light scattering, in contrast, gives an absolute molecular mass that is independent of the shape of the molecule. We therefore carried out on-line light scattering experiments in which the purified enzyme was loaded on a Superose 12 size exclusion column and the eluant monitored.
by laser light scattering. This technique gave a molecular weight of 179 kDa for M.AhdI (Fig. 2b).

**Analytical centrifugation**

Analytical ultracentrifugation was used to further characterise the behaviour of the multisubunit enzyme. Sedimentation equilibrium can yield information on concentration-dependant oligomerisation. M.AhdI was therefore sedimented to equilibrium over the concentration range 0.6–19.4 μM, using three different rotor speeds, and the data analysed globally using the least squares fitting program Nonlin (v. 1.06). The data could not be fitted to a single species model or to a two component independent species model. However, a good fit was obtained using a dimer–tetramer equilibrium model, with molecular weights of 85 and 170 kDa, respectively, from which the dissociation constant $K_d = 2.2 \mu M$ was obtained (Fig. 3).

Sedimentation velocity experiments were then carried out on the purified MTase. From analysis of the sedimentation profiles using the DCDT algorithm (15) two components were observed, a major component with $S_{20,w} = 7.8 S$ and a minor component with $S_{20,w} = 4.6 S$. Assuming the faster sedimenting species corresponds to M$_2$S$_2$ and the slower one to MS, using the molecular weights of these species one can calculate frictional ratios ($f/f_o$) of 1.34 and 1.43, respectively. This

---

**Figure 3.** Sedimentation equilibrium. The top graph shows the sedimentation equilibrium data for the methyltransferase at a series of different concentrations and rotor speeds. Each data set has been fitted with a global model to describe the data and is shown by the black line, with the residuals to the fit plotted below. The model is of two heterodimers of 85 kDa self-associating to form a tetramer of 170 kDa.
indicates that both species are somewhat elongated, although the tetramer is more compact than the dimer, suggesting that dimers associate in a side-by-side manner, rather than end-to-end, when forming the tetramer.

DNA binding studies

Gel retardation assays were carried out using a 30 bp oligonucleotide containing the AhdI DNA recognition site, GACN5GTC. The M.AhdI–DNA binding interaction was optimised for binding buffer conditions, varying the pH from 6.8 to 9.0, the salt concentration from 0 to 200 mM NaCl and with the addition of various concentrations of MgCl₂, ATP and AdoMet. The variety of conditions tested resulted in no discernable differences in the gel retardation patterns between the different binding buffers or the presence of MgCl₂, ATP and AdoMet (data not shown).

Titration of M.AhdI with the specific 30 bp oligonucleotide shows the presence of up to three complexes. At the lower protein concentrations, there is a clear complex band and a much fainter diffuse band (Fig. 5a, bands 2 and 3, respectively). These most likely correspond to binding the M₂S₂ tetramer and the MS heterodimer, respectively, and are only seen with the specific duplex (see below). At higher concentrations of protein, there is a clear complex band and a much fainter diffuse band (Fig. 5a, bands 2 and 3, respectively). These most likely correspond to binding the M₂S₂ tetramer and the MS heterodimer, respectively, and are only seen with the specific duplex (see below). At higher concentrations of protein, there is a clear complex band and a much fainter diffuse band (Fig. 5a, bands 2 and 3, respectively).

Gel retardation assays were also carried out using 20, 30 and 60 bp DNA duplexes bearing the AhdI recognition site. Similar binding characteristics were seen for the 30 and 60 bp duplexes, but much weaker binding to the 20 bp DNA was observed resulting in substantial smearing down the track (data not shown). This indicates that there are interactions outside the 11 bp recognition site that stabilise the protein–DNA complex. Further DNA duplexes, with alterations to the recognition site, were used to check the specificity of the protein–DNA interaction. Binding to a 30 bp duplex lacking the recognition site (Fig. 5b) resulted in a large reduction in binding affinity so that specific complex formation (band 2) was not apparent, although the slow migrating complex (band 1) was present at the higher protein concentrations, which we attribute to non-specific binding of multiple copies of the enzyme. An identical result was seen using a 30 bp duplex bearing one half-site from the AhdI recognition sequence. Gel retardation assays using a non-specific 30 bp duplex in competition with the 30 bp duplex containing the recognition sequence showed that the former did not effectively compete for binding, even when present in a 25:1 excess (data not shown).

Quantitation of the gel retardation assay using the 30 bp duplex containing the recognition sequence allowed the binding curve for M.AhdI to be fitted (Fig. 5c). The curve is
sigmoidal, indicative of cooperative binding, and could not be fitted by the standard single site binding model. This would be expected if the active binding species is the tetramer, since this will dissociate at low concentrations. The band that we suggest arises from the MS heterodimer binding to DNA is extremely weak, and to a good approximation this interaction can be ignored. Using a cooperative binding model which allows for the formation of M2S2 from MS dimers (with dissociation constant $K_1$) and assuming only the tetrameric species is capable of binding to DNA (with dissociation constant $K_2$), the best fit was obtained with $K_1 = 1 \mu\text{M}$ and $K_2 = 50 \text{nM}$.

**Enzyme activity**

An enzyme assay for M.AhdI was devised, based on the ability of the MTase to methylate the AhdI recognition sequence and thus prevent cleavage by the cognate endonuclease, R.AhdI. The extent of methylation at a given time can be assessed by the inhibition of endonuclease activity, which in the absence of methylation results in the cleavage of the linearised 2686 bp plasmid into two fragments of almost equal length (1298 and 1388 bp), which can be visualised by agarose gel electrophoresis (Fig. 6a). Further experiments showed that the methylation reaction is dependent on the presence of both AdoMet and Mg$^{2+}$ ions, and is inhibited by EDTA. Figure 6b shows a time course of the methylation reaction. This curve can be fitted using the integrated form of the Michaelis–Menten equation, from which the values for $K_m$ for the DNA substrate (1.4 nM) and $V_{\max}$ (0.25 nM min$^{-1}$) were estimated, the latter providing an estimate of $k_{\text{cat}}$ ($4.2 \times 10^{-3}$ s$^{-1}$). However, these must be regarded as provisional values, pending a more detailed enzyme kinetic analysis.

**DISCUSSION**

The results of the analytical centrifugation experiments show that M.AhdI exists as a 170 kDa tetrameric enzyme (M$_2$S$_2$) which dissociates into two heterodimers of the form MS at low concentration. Gel retardation analysis indicates that only the tetrameric form binds to DNA with high affinity. Thus in the intact enzyme, two copies of the AhdI S subunit combine with two M subunits (Fig. 7) to form a symmetrical tetrameric methyltransferase (M$_2$S$_2$), the naturally occurring equivalent...
of the structure proposed for deletion mutants of classical type I enzymes (16–18). The organisation of the subunits of M.AhdII is thus comparable with that resulting from a combination of the native M subunit of the EcoR124I R-M system with the N-terminal half of the S subunit of EcoR124I (denoted S3) and shows similar characteristics to the engineered enzyme. The (S3/M)2 enzyme also methylates a symmetrical recognition sequence (in this case GAAN7TTC) and dissociates at low concentration, resulting in a sigmoidal DNA binding curve resembling that observed for M.AhdII (19).

The \( K_d \) for DNA binding of M.AhdII (50 nM) is a little higher than that reported for the (S3/M)2 complex (\( K_d = 5 \text{ nM} \)) and other type I enzymes such as EcoKI and EcoR124I (\( K_d = 10 \text{ nM} \)), but is nevertheless at the lower end of the range of \( K_d \) values typical for MTases binding sequence-specifically to oligonucleotide duplexes (\( 10^{-9} - 10^{-8} \text{ M} \)). It should be borne in mind, however, that dissociation of the tetrameric M.AhdII MTase into heterodimers means in practice that the effective DNA binding affinity is substantially lower than would be predicted from the \( K_d \) for the tetramer–DNA interaction.

The \( K_m \) (for its DNA substrate) of M.AhdII (1.4 nM) is much lower than that reported for the type I MTase M.EcoKI (880 nM) (1) and is closer to the \( K_m \) reported for the type II MTase M.EcoRI (0.3–1.3 nM) (20,21). Likewise the value of the catalytic rate constant, \( k_{cat} \), for M.AhdII (4.2 \( \times 10^{-3} \text{ s}^{-1} \)) is intermediate between that of M.EcoRI (0.014–0.053 \( \text{ s}^{-1} \)) (20,21) and M.EcoKI (2 \( \times 10^{-4} \text{ s}^{-1} \)). Caution should be exercised in interpreting these results, since the enzyme kinetic experiments are often conducted under different solution conditions and using different length DNA substrates (varying from short oligonucleotide duplexes to plasmids). It will also be apparent that if dissociation of the enzyme affects DNA binding, it should reduce the rate of DNA methylation at low enzyme concentrations. Since the enzyme assay is performed at a concentration (1 nM) well below the \( K_d \) (\( \sim 1 \mu \text{ M} \)) for dissociation of the enzyme, this is likely to affect the activity of M.AhdII. In the bacterial cell, it has been estimated that MTase levels are likely to be in the region of 50–500 nM (1) and thus dissociation of M.AhdII is likely to occur in the cell, unless macromolecular exclusion effects prevail.

\( \text{AhdII} \) may represent an evolutionary intermediate between type I and type II R-M systems. Indeed, searching the sequence database (using BLASTP) shows that there is a significant homology between the M subunit of \( \text{AhdII} \) and the type II methyltransferase \( \text{XmnII} \) (as well as to several type I systems); a similar search with the S subunit of \( \text{AhdII} \) shows significant similarity only to the S subunits of type I enzymes. Even so, in both cases the homology is rather limited.

The S subunit of \( \text{AhdII} \) has a short sequence of \( \sim 20 \) amino acids (amino acids 168–187) that is clearly related to the sequence found at the start of the central conserved region of all type I S subunits (Fig. 7a), with 63% identity (79% similarity) to the consensus sequence at this site. However, there are no extensive sequence similarities of the sort commonly found between members of the same family. The S subunit of \( \text{AhdII} \) would appear to comprise a single N-terminal TRD (approximately amino acids 1–167) responsible for recognition of the trinucleotide half-site GAC and a single domain (approximately amino acids 168–227) that is equivalent to the ‘conserved’ domain of type I enzymes and is therefore presumed to interact with the M subunit (22). These features in the sequence of the S subunit of \( \text{AhdII} \) together with the symmetrical DNA recognition sequence and the observed stoichiometry of the intact \( \text{AhdII} \) Mtase, suggest that the \( \text{AhdII} \) S subunit is equivalent to one half of a type I specificity subunit (Fig. 7b and c).

It is likely that \( \text{AhdII} \) represents a primitive type I Mtase having a single TRD, from which more complex type I enzymes evolved by gene duplication and fusion. Further evolution of different specificities may then have arisen by ‘shuffling’ of TRDs on a scaffold of ‘conserved’ regions, a process that can also be engineered in vitro (11,23).

ACKNOWLEDGEMENTS

We thank Dr Darren Mernagh for advice and discussion during the earlier stages of this project and Polymer Laboratories for generous provision of laser light scattering facilities. The work was supported by a research grant from the Wellcome Trust and a PhD studentship from the Institute of Biomedical and Biomolecular Sciences, University of Portsmouth.

REFERENCES


